Open Data as a New Commons

Empowering citizens to make meaningful use of a new resource

Morelli, Nicola; Mulder, Ingrid; Concilio, Grazia; Pedersen, Janice; Jaskiewicz, Tomasz; Götzen, Amalia De; Aguilar, Marc

Published in:
Internet Science

DOI (link to publication from Publisher):
10.1007/978-3-319-70284-1_17

Publication date:
2017

Document Version
Accepted author manuscript, peer reviewed version

Open Data as a New Commons. Empowering citizens to make meaningful use of a new resource

Nicola Morelli[1], Ingrid Mulder[2], Grazia Concilio[3], Janice S. Pedersen[4], Tomasz Jaskiewicz[2] , Amalia de Götzen[1] , Marc Arguillar[5]

[1] Aalborg University, Copenhagen, Denmark
[2]Technische Universiteit, Delft, Nederlands
[3] Politecnico di Milano, Milano, Italy
[4] Antropologerne ApS, Copenhagen, Denmark
[5] I2CAT, Barcelona, Spain
nmor@create.aau.dk

Abstract. An increasing computing capability is raising the opportunities to use a large amount of publicly available data for creating new applications and a new generation of public services.

But while it is easy to find some early examples of services concerning control systems (e.g. traffic, meteo, telecommunication) and commercial applications (e.g. profiling systems), few examples are instead available about the use of data as a new resource for empowering citizens, i.e. supporting citizens' decisions about everyday life, political choices, organization of their movements, information about social, cultural and environmental opportunities around them and government choices. Developing spaces for enabling citizens to harness the opportunities coming from the use of this new resource, offers thus a substantial promise of social innovation.

This means that open data is virtually a new resource that could become a new commons with the engagement of interested and active communities. The condition for open data becoming a new commons is that citizens become aware of the potential of this resource, that they use it for creating new services and that new practices and infrastructures are defined, that would support the use of such resource.

Keywords: Open Data, Social Innovation, commons.

1. Introduction

The term commons was used in medieval England to indicate parts of land which were openly accessible to everybody in a community. By extension the term was lately used to indicate any resource (land, forests, water basins, or pastures) that was freely available but controlled by a community, through a set of rules that regulate the access in order to create a sustainable use of the resource.

The open availability of such resources is of course not a sufficient condition for them to survive as a commons and be used in a sustainable way. Without a community taking care of a commons pasture, for instance, this land would be neglected or someone would claim its ownership. In order for a commons to persist and grow, it is necessary that a defined group of people develop their own distinctive social practices and bodies of knowledge for managing the resource (Ostrom, 1990).

This is also true for "modern" resources, or "new commons" (Hess, 2008), of which a very famous example is the widely used open software platform known as GNU/Linux (Bollier, 2014). If a commons can be defined as *a resource + a community + a set of social protocols* (Bollier, 2014, p. 15), what is still to be designed in order to define open data as a new commons? Why it is important do definit it as such? What are the challenges and the opportunities in the current socio-technical systems?

The latest development of ICT, Internet of things and communication technologies are making it possible to collect a very large amount of data. Computing capabilities are in turn making it possible to coordinate the collection of those data in coherent data sets that can provide information on many aspects of our life, our cities, our environment, social life and healthcare condition. This increases our capability of monitoring our environment.

The potential of this phenomenon is still largely unexplored, because if, on one hand, the input side is becoming very diversified and providing rich data sets, on the other hand the output side is only used by few stakeholders, that have identified good opportunities for their business. Many businesses and institutions take advantage of large and typically closed data sets. For example, Google's core business is to collect data about users of its services such that it can better target its advertisements to individuals; insurance companies collect data about their customers to assess insurance risks; and tax authorities use big data to identify cases of tax fraud. On the other hand, there are many data sets that do not contain sensitive information and that already are, or could be, made open by governments and companies for public use. However, non-experts do not have the knowledge and skills needed to take advantage of such large data sets in the way that large institutions or companies do.

1.1. The input side

An increasing presence of microprocessors, sensors and similar devices to control several aspects of urban and business life, such as traffic, access to public buildings or environmental conditions is generating a large amount of data that can be easily collected and made available. Without touching the delicate issue of privacy, related to this data collection, which opens ethical questions that cannot be ignored, the availability of those data could also be and in some case is being used as an opportunity to control crucial aspects of our environments, from parking to traffic regulation, environmental monitoring and garbage collection.

The miniaturization of all this technology made it possible to have wearable devices that can also be used to consciously track ourselves: we can monitor our body in many different ways, record how many steps we have taken in one day and so on. This is what is called the quantified self, a trend in continuous expansion. Such trends deal with very private information, but we often decide to share those data among friends or even publicly, adding information that also others can use (e.g. on weather conditions, condition of the area where we have just been running, etc).

Another important source of data comes from social networks, in which users are more or less voluntarily providing data about their preferences, geographical positioning, political orientation, and even private life. Those data have so far been used for

commercial purposes, because they give chances for accurate profiling of social network users and therefore make it possible to generate highly personalized commercial offers. However social networks have also been used for different purposes; they have been working as an aggregation tool for movements, local initiatives, or political groups. Social network users often provide on-time qualitative data about places, weather, transport problems, healthcare crisis, emergency, security issues, catastrophes,etc.

The integration of those two kinds of datasets, the automated data input coming from micro-controllers disseminated in our daily environments and the voluntary data input provided by users could improve our understanding of the context in which we live in, because it would link quantitative data, often hard to interpret, to qualitative data, which are sometimes imprecise, but very useful to tell a narrative, allowing us to better navigate in such a complex context, giving meaning to numbers. This paper argues that such integration presents opportunities to consider data as a new public resource. For this resource to be considered as a commons a new awareness should be raised on the output side, to make sure that more people – citizens, organisations, interest groups, public administrations – be interested in using this resource and developing tools and shared practices to take care of this resource

1.2. The output side

The initiatives to make datasets available to the broader public are becoming more and more frequent. Several cities, such as Copenhagen, Rotterdam, Glasgow, Barcelona etc. have opened their databases, several datasets are now freely accessible on the Web and can be used by anyone.

The question is how many stakeholders have the capability to use those data for relevant applications. The use of those data in fact requires:

1. technical skills, to create applications based on the available data,
2. creative capabilities, to understand which needs those datasets can address and for which part of our society,
3. data literacy, to pose answerable research questions and interrogate the available data (or create a dataset).

It has already been mentioned that some early examples of the use of the datasets concern control systems (e.g. traffic, meteo, telecommunication) and commercial applications (e.g. profiling systems); few examples are available about the use of such data as a new resource for empowering citizens, i.e. supporting citizens' decisions about everyday life, political choices, organization of their movements, information about social, cultural and environmental opportunities around them, government choices. Developing means for enabling citizens to harness the opportunities coming from the use of this new resource, offers thus a substantial promise of social innovation.

2. Context: the Open4Citizens project

Looking back at open data, although this resource has been seized by many companies to be used for commercial purposes, its potential is still widely unexplored and could still have large scope for applications and uses that can directly benefit communities, even outside the market mechanisms. Open data can be the base of a new generation of public services that is directly defined, designed and used by citizens. Open data can become a commons if communities are built around them, that create sets of practices and rules to use this resource in the most sustainable way. This is the aim of the EU-Funded Open4Citizens project (www.open4citizens.eu), which is experimenting on new forms of collaboration between citizens, public authorities, interest groups, local businesses and IT experts, with the aim of a) generating new services, b) aggregating communities around the resource c) creating new infrastructures for the use of open data.

The project focuses on Open Data as a new shared resources and aims at generating the conditions for this resources to be used and managed as a commons. More specifically the project is developing strategies and tool with the aim of:

- Aggregating a community that takes care and uses the resource
- Developing a set of practices for using the resource
- Infrastructuring the community with a set of tools that facilitate the access to the resource and its usage

The project is articulated in 5 pilots in Barcelona, Copenhagen, Karlstad, Milano and Rotterdam. Each of these pilots works on different challenges ((Table 1) to define different solutions, within a shared framework of design processes, tools and methods to increase citizens' awareness of open data and engage them in the creation of new solutions for their everyday problems.

This paper focuses on the creation of the needed infrastructure for the exploitation of open data as a new commons, and on the aggregation of communities (ecosystems) of users highlighting the hackathon process as an instrumental mean to facilitate the creation of this culture.

2.1. Aggregating communities

Citizens' participation in the Open4Citizens pilots and their awareness in the use of open data are supported through the organization of *hackathons*.

Hackathons are a common practice among information technology (IT) enthusiasts and there is an extensive literature reporting on hackathon experiences (Linnel, 2014; Tanenbaum, 2014; Hecht, 2014). They consist of a 2-3 day *pressure cooker* event in which groups of IT experts collaborate to develop new applications. They are a playground for IT experts, that can propose new solutions, usually very innovative from a technological point of view. However these solutions rarely address citizens' needs, because of the nature of the hackathon, which is gathering experts coders, neglecting though the prospective users of such solutions/apps (i.e. those who are able to propose meaningful problems).

The purpose of the Open4Citizens hackathons is to extend the participation to non-IT experts, that means common citizens, interest groups, public servants and business

organisations, which can together work on a commonly defined problem area. For this reason the phase of organization of each hackathon is critical, to ensure the participation of people with different skills, motivations and knowledge. Each Open4Citizens hackathon aspires to put together a local community that shares values, interests and motivations in specific problematic areas, thus creating an *ecosystem* to generate or consolidate the demand for open data or to organize crowdsourced gathering of new open data. In order for a common good to benefit members of a community, there is a need for a shared understanding of the value of this resource as well as clear guidelines for its use. The Open4Citizens hackathon approach focuses on placing ordinary but engaged citizens at the centre of the development of open-data driven solutions to societal challenges that are important to them. This approach ensures that collaboration between hackathon participants with different skill sets and interests is appropriately facilitated such that there is a shared understanding of the potential of open data, all participants have the opportunity to manipulate open data and the contributions of all participants to the co-creation of solutions are valued. Some of the main competences and skills for the hackathons in Open4Citizens are listed in Table 2.

Table 1: Competences and skills in the ecosystem of stakeholders participating in the hackathon

Stakeholder	**Knowledge/Experience/Skills**
Citizens	Personal/Daily life experience
Interest groups	Domain experts, issue experts
IT experts/hackers	Technical knowledge
Public authorities	Institutional knowledge (regulations, public policies)
Data experts	Data literacy (scraping, cleaning, visualizing)
Entrepreneurs	Business knowledge
Data Owners	Data format, potential of existing data, available data

2.2. Building practices: the hackathon process

The hackathon approach is meant to generate a set of practices to support citizens participation. Unlike the better known hackathons, the O4C hackathons aggregate quite heterogeneous knowledges and cultures; therefore the process needs to be accurately designed to reconcile dissimilar perspectives. The O4C team organized the hackathon as a process including:

- An initial **pre-hack** phase, characterized by an intense activity of dataset identification, verification and specification, running in parallel to the definition of the challenge(s) that the hackathon is supposed to tackle.

- The **hackathon event,** a *pressure-cooker* co-design event involving all the relevant stakeholders; and
- A **post-hack** phase, including follow-up activities to implement technical and business aspects of the outcomes of the post-hack.

2.3. Infrastructuring for using open data

The hackathon starter kit

In order to support participation in a hackathon by anyone with an interest in open data or in the societal challenges to be solved, a starter kit is proposed, that includes tools (Figure 1) that support citizens in the various steps in a hackathon process while guiding them progressively from inspired idea to concrete solutions.

Each hackathon process involves an interplay between three types of activities, which are: 1) "inspiration", understanding the opportunities on the intersection of citizens' needs and open data current or potential availability, 2) "ideation", i.e., creatively addressing these opportunities, and 3) "implementation", translating the ideas into concrete app or service concepts and prototypes. Physical (paper) tools complement and interact with the Open4Citizens platform (www.opendatalab.eu), that also includes visualization development tools.

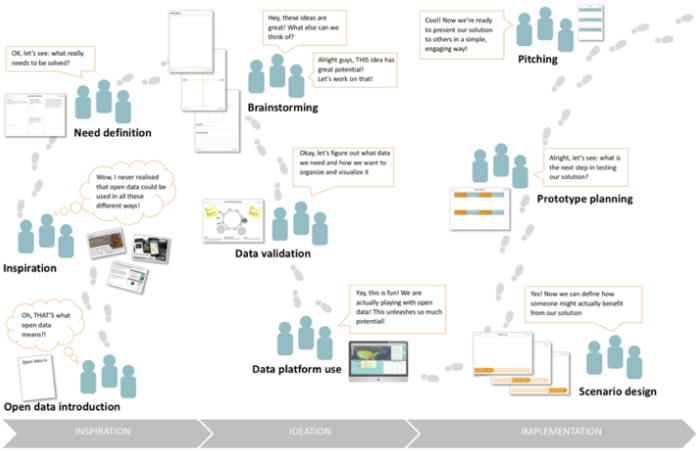

Fig. 1: Hackathon starter kit

The Open4Citizens (O4C) Platform

The O4C platform is a digital service for sharing, aggregating, curating and supporting open data and data-related knowledge, know-how and best practices within and across engaged communities. The platform serves as a backbone underlying the hackathons in the five pilot cities involved at this stage, supporting a shared understanding across

different societal contexts and thematic challenges to be addressed using open data. It is intended to have low skill entry point and be user friendly so that citizens, hackathon organisers, facilitators and other stakeholders are able to use it, even if they have no previous experience of working with open data. It allows for shared open data usage across locations and for feedback to be gathered regarding its usability.

The O4C platform is being adapted in line with learning from hackathons to support the development of a network of pilot OpenDataLabs in the five project locations: Barcelona, Copenhagen, Karlstad, Milan and Rotterdam. Pilot hackathons and related activities provide test cases for the development of a shared vision of OpenDataLabs including the role that the O4C platform should play in underpinning the lab network.

The OpenDataLab

OpenDataLab represents the main conceptual and operational output of the Open4Citizens project. In the ongoing project an initial OpenDataLab has been conceptualized, drawing inspiration from FabLabs, where citizens' good ideas about new products can find knowledge and technologies to support the realisation of their ideas. Likewise, in OpenDataLabs citizens, groups or associations will be able to develop their ideas for new services using existing open data. The Open4Citizens project has already produced some deeper understanding of what an OpenDataLab can be. Surely it represents the space where an ecosystem of citizens, ICT experts and Public service managers is activated to generate solutions to urban problems by using open data. Furthermore the OpenDataLabs will be supported by the online platform presented in the previous section.

3. The first hackathons' cycle

As already mentioned, the O4C project is organized in 5 pilots where the above mentioned tools and methods developed by the consortium are tested and evaluated in two hackathon cycles. The evaluation of the first cycle has been the building blok for the design of the second one, that will run in fall 2017.

The experiments on open data in each pilot had to be articulated along three main dimensions: 1) the availability of data, 2) the nature of the societal challenge to be addressed and 3) the creation of a community, i.e. people with different interests, motivations, skills and knowledge that could experiment on practices to work with open data.

Those dimensions have been important to define the thematic areas for the activities in each pilot, which are synthesized in Table 2.

Table 2: Challenges in each Open4Citizens pilot

Pilot	Thematic areas
Copenhagen (DK)	Migration, Integration
Karlstad (SE)	Healthcare, quantify self

Rotterdam (NL)	Parks and common spaces
Barcelona (SP)	Healthcare, Culture
Milano (IT)	Transparency in public decisions for urban transformation

If we focus on the outcome of the hackathons, different results can be highlighted on the basis of the variation of the three above mentioned dimensions: data, challenges and people. The proposed solutions ranges from apps, public services, visualisations, online plaforms, physical places, online portals etc. and therefore also their level of development varies from concepts to prototypes, that are being developed as new services or as part of existing services.

As examples, *Seek a seeker* is a platform to connect potential hosts and asylum seekers. Locals (Danes and former asylum seekers) "adopt" or host a new asylum seeker for a period of time. The platform uses open data in a crowdsourced way. *Cantieri miei* is a service that uses existing open data to help citizens affected by the metro construction works in Milano in getting compensation payments. *PolenCat* is an app (under development) which integrates an open database of historical and current pollen levels with the user-contributed medical and geo-localisation data, and recommends personal daily routes that avoid or mitigate those areas expected to have high concentrations of pollens and other allergens to which the user is susceptible (see fig. 1).

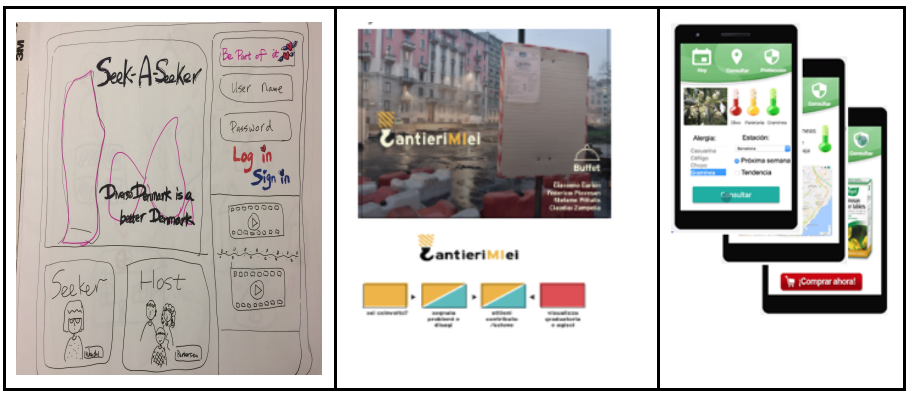

Figure 1: Three solutions: Seek a seeker - the concept of a platform from Copenhagen; CantieriMiei - a service under develpement from Milano; PolenCat, an app under development from Barcelona.

In these three examples the use of open data has also varied a lot: from imagining crowd sourced data sets in the Copenhagen pilot to a specific data set that was open from the municipality of Milan explicitely for the hackathon, to the use of available open datasets in the Barcelona's pilot. The question of transparency in Milan, in fact, was directly referring to the availability of data to describe specific problems within the city. In that case the main effort has been focused on gathering a number of relevant

stakeholders, which own the data and could suggest meaningful uses of them. In other pilots, where the problem was more widely defined (e.g. immigration/integration in Copenhagen) the proposed challenge did not directly point of relevant datasets, and the work of the participants has consisted in defining possible solutions on the basis of datasets to be crowdsourced.

All in all it is clear that significant work is needed in the pre-hack phase and into the hackathon itself to make data available in relevant formats and to provide guidance on how to use it. In this regard, the OpenDataLabs could have a crucial role and the first hackathon cycle has been a test-bed for developing different possible OpenDataLabs and identifying how they can become a sustainable, linked network.

4. Discussion

It could be argued that the current state of play regarding useful access to and use of open data by anyone wishing to do so is more reminiscent of the old 'tragedy of the commons' (Hardin 2009) idea than we would like. This can more appropriately be defined as the 'tragedy of open access regimes' (Daly & Farley 2011) when relating to empowering citizens to make appropriate use of open data. In this understanding, a resource that should be available to anyone can in fact only be used by a small subset of individuals or organisations with the skills and capacity to benefit from its potentials. In the case of open data, the Open4Citizens project suggests that the skills needed are technical skills, e.g. for creating apps or other means of open data integration in solutions, and creative capabilities, e.g. capabilities to identify opportunities where open data could be useful and imagining how they might be used.

The project team is testing how best to make these skills available to citizens or to support them in gaining them. By developing fit-for-purpose tools, guidelines for how to use them, and a support infrastructure for creating potentially impactful solutions to real-world challenges in urban service delivery, Open4Citizens seek to empower citizens to use the open data commons.

It has been demonstrated that understanding local successes in sustainable and equitable use of common resources provides lessons for use of large-scale commons (Ostrom et al 1999). By learning from and comparing the facilitators and barriers relating to open data use for the common good in five different urban contexts across Europe, the project is laying the basis for an understanding of how an open data movement might be achieved. This movement is defined by the use of open data by any individual or community wishing to do so in a way that potentially benefits all citizens.

In order to learn from successes and challenges in open data use for the common good, Open4Citizens is working towards understanding the complex interactions at the local pilot level between 1) people interested in the potentials and practical applications of open data, 2) the nature of the open data landscape locally, i.e. how much data is open, how open it is in practice, whether the available open data is also useful and 3) the types of societal challenges that can appropriately be addressed using open data.

Beside the concrete solutions the Open4Citizens hackathons lead to, the project is also emphasising the potential of a hackathon event for fostering communities and

open-data related knowledge, skills and know-how in these communities. Consequently, the OpenDataLab is a means to sustain and nurture these communities beyond hackathon events which are limited in duration.

5. Conclusions

The use of technology by human beings is always mediated (Miller & Slater 2000).

Technological elements such as open data, data manipulation tools and the fit-for-purpose OpenDataLab platform are only one, albeit essential, piece to establish an open data movement where open data truly become a common good. To this end, it is important to understand the needs of individuals and communities wanting to make use of this commons as well as the barriers and opportunities related to their open data use. It is also very important to lower the barriers to participation, so that contributing to this work isn't dependent on mastery of information technologies. For this reason, the knowledge base and guidelines being developed in this project regarding the interactions between analog tools and the digital, technological tools are key.

A central element in applying emerging lessons from local pilots to a larger scale European or even global open data movement involves understanding the nature of the commons: both the open data as a resource and the community or communities that want to use it. Linking these requires a shared understanding by the entire community (citizens, academics, the public sector and the private sector) of how open data can be used and to what end(s).

References

Bollier, D. (2014). *Think like a commoner : a short introduction to the life of the commons.* Gabriola Island: New Society Publishers.

Daly, H. E., & Farley, J. (2011). *Ecological economics: principles and applications*. Island press.

Hardin, G. (2009). *The Tragedy of the Commons*, Journal of Natural Resources Policy Research, 1(3), 243-253.

Hecht, B. A., Werner, J., Raskar, R., Jouttenus, T. T., Khandbahale, S. S., Jouttenus, M., & Bell, P. (2014, March). *The KumbhThon technical hackathon for Nashik: A model for STEM education and social entrepreneurship.* In Integrated STEM Edu cation Conference (ISEC), 2014 IEEE (pp. 1-5). IEEE.

Hess, C. (2008) *Mapping the new commons*, The twelfth Biennial Conference of the international association for the study of the Commons, Cheltenham, UK, 24-28 July.

Linnell, N., Figueira, S., Chintala, N., Falzarano, L., & Ciancio, V. (2014, October). *Hack for the homeless: A humanitarian technology hackathon*. In Global Humani tarian Technology Conference (GHTC), 2014 IEEE (pp. 577-584). IEEE.

Miller, D., & Slater, D. (2000). *The Internet: an ethnographic approach*. Berg.

Ostrom, Elinor. 1990. *Governing the Commons. The Evolution of Institutions for Col lective Action*. Cambridge: Cambridge University Press.

Ostrom, E., Burger, J., Field, C. B., Norgaard, R. B., & Policansky, D. (1999). *Revisit ing the commons: local lessons, global challenges*. science, 284(5412), 278-282.

Tanenbaum, K., Tanenbaum, J. G., Williams, A. M., Ratto, M., Resch, G., & Gamba Bari, A. (2014, April). *Critical making hackathon: situated hacking, surveillance and big data proposal*. In CHI'14 Extended Abstracts on Human Factors in Compu ting Systems (pp. 17-20). ACM.